OPEN ACCESS

Indonesian Journal of Health Research and Development Vol 3 No 3 (2025) https://doi.org/10.58723/ijhrd.v3i3.507

Check for updates

Research Article

Deep Learning-Based ResNet-50 Transfer Learning Approaches for Pneumonia Detection from Chest X-Ray Images: With and Without Fine-Tuning

Received: August 27, 2025	Revised: November 14, 2025	Accepted: 30 November 2025	Published: 30 November 2025

Dwi Robiul Rochmawati*, Lidya Maryani

Abstract

Background: Pneumonia remains one of the leading causes of morbidity and mortality worldwide, particularly among children and older adults in low-resource settings. Diagnosis based on chest X-ray interpretation often depends on radiologist expertise, which may be limited in availability and prone to subjectivity. Deep learning offers a promising alternative to improve diagnostic efficiency and consistency.
Aims: This study aims to evaluate the effectiveness of the ResNet-50 architecture for pneumonia detection using chest X-ray images by comparing transfer learning with frozen layers and partial fine-tuning strategies.
Methods: A total of 5,856 chest X-ray images were obtained from a public dataset and divided into training, validation, and testing sets using stratified sampling. Data preprocessing included resizing, normalization, and augmentation. Two models were developed: (1) a frozen ResNet-50 model, where all convolutional layers were fixed, and (2) a fine-tuned ResNet-50 model, where the final convolutional layers were retrained. Performance was evaluated using accuracy, precision, recall, F1-score, and area under the ROC curve (AUC). Statistical tests were conducted to assess performance differences between the two models.
Results: The frozen model achieved an accuracy of 62.50% and an AUC of 0.4819, indicating weak classification performance. In contrast, the fine-tuned model demonstrated substantially higher accuracy of 85.90%, F1-score of 0.8967, and AUC of 0.9510, showing strong discriminative capability. Statistical analysis confirmed that the performance improvement in accuracy was significant.
Conclusion: Fine-tuning significantly enhances the applicability of ResNet-50 for pneumonia detection. Without feature adaptation, pretrained models struggle to generalize to medical imaging domains. Fine-tuned transfer learning provides a more reliable framework for developing computer-aided diagnostic systems, particularly in clinical environments with limited expert availability.

Keywords: Chest X-ray; Deep learning; Transfer learning; Pneumonia detection; ResNet-50

1. INTRODUCTION

Pneumonia is one of the deadliest respiratory diseases globally, particularly in developing countries. According to the World Health Organization (WHO), pneumonia causes more than 2.5 million deaths each year, with high prevalence among children under five and the elderly (Sharma & Guleria, 2024; Surani & Jain, 2025). This disease is caused by various pathogens such as bacteria, viruses, and fungi, which induce inflammation in the pulmonary alveoli (Virmani et al., 2023).

Early and accurate detection of pneumonia is crucial for effective treatment. Traditional methods that rely on manual interpretation of chest X-ray images are often slow and prone to human error, especially in resource-limited regions (Aditya Pai et al., 2024; Munna et al., 2025), and although chest X-ray imaging is the primary modality for pneumonia diagnosis, its effectiveness is hampered by variability in image interpretation, dataset imbalance, and resource constraints, especially in rural and underdeveloped regions (Dabre et al., 2024; Rani & Kaur, 2024). Recent advances in artificial intelligence (AI), particularly deep learning and transfer learning, offer promising avenues to address these challenges. Among various architectures, ResNet-50 has emerged as

Publisher Note:

CV Media Inti Teknologi stays neutral about jurisdictional claims in published maps and institutional affiliations.

Copyright

©2025 by the author(s).

Licensee CV Media Inti Teknologi, Bengkulu, Indonesia. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-ShareAlike (CCBY-SA) license (https://creativecommons.org/licenses/by-sa/4.0/).

a robust backbone for medical image analysis, with transfer learning and fine-tuning strategies demonstrating substantial improvements in diagnostic accuracy and efficiency (Bhola & Gulhane, 2024; Khojare & Bhattacharjee, 2025; Maheshwar et al., 2025; Sohel et al., 2024; Vemuri et al., 2025).

Although deep learning technology has shown great potential, several challenges remain. Major challenges include difficulty in identifying abnormal regions due to interference from other tissues, as well as pneumonia features that frequently overlap with other diagnoses (Wang et al., 2019). In addition, imbalanced datasets and scarcity of well-labeled data also pose significant obstacles (Alshanketi et al., 2025).

Several approaches have been attempted to address these issues, including the use of transfer learning techniques and hybrid models. For example, the YOLOv6 model has demonstrated outstanding performance in real-time pneumonia classification (Munna et al., 2025). Furthermore, preprocessing techniques such as histogram equalization and Gaussian filtering have been utilized to enhance feature extraction (Munna et al., 2025; Shirvalkar & Remya Ajai, 2024).

Previous studies have extensively explored various deep learning architectures for pneumonia detection using chest X-ray images, with VGG16, ResNet50, and DenseNet121 being the most frequently investigated models. VGG16 demonstrates high accuracy, achieving 90.87% in one study and 96.6% with 98.1% sensitivity and 97.2% precision in another, while hybrid models combining VGG16 and DenseNet121 reach 92.53% accuracy (Shah et al., 2020a; Shrimali, 2024). ResNet50 also shows strong performance with 89.06% accuracy and 86.23% recall, alongside a combined model accuracy of 92.03%, and it often outperforms other models in accuracy and recall (Gupta et al., 2025; Manickam et al., 2021; Michele & Kusuma, 2021; Yaqin et al., 2024). DenseNet121 offers balanced metrics, achieving an F1-score of 74.52% and a high validation accuracy of 98.68% (Shakeri & Far, 2024a; Shrimali, 2024), competing closely with ResNet50. Comparatively, VGG16 and ResNet50 are known for their high accuracy, while DenseNet121 provides stable and reliable performance (Shah et al., 2020b; Shakeri & Far, 2024b; Shrimali, 2024). Hybrid and ensemble approaches, including combinations of VGG16 and DenseNet121, further enhance accuracy and recall (Paul & Naskar, 2023; Shrimali, 2024). Additional architectures such as MobileNetV2 and EfficientNetB0 show competitive performance (Shrimali, 2024), while more recent models like Vision Transformers (ViT) are being explored for their potential to handle larger datasets and improve AI explainability (Alshanketi et al., 2025; Jihad et al., 2025). The exploration of deep learning architectures for pneumonia detection reveals that models like VGG16, ResNet50, and DenseNet121 are highly effective, each with unique strengths. Hybrid and ensemble approach further enhance diagnostic accuracy, making these models valuable tools in clinical settings for timely and accurate pneumonia detection.

Despite previous research demonstrating the effectiveness of various deep learning architectures, including VGG16, ResNet50, and DenseNet121, in pneumonia detection, there are still several gaps that need to be addressed. First, while ResNet50 has shown strong performance, there is limited research on the application of transfer learning with and without fine-tuning for this model. Additionally, many studies do not directly compare the impact of fine-tuning on accuracy and other performance metrics in the context of pneumonia detection from chest X-ray images. Existing research also often fails to explore the potential of ResNet50 in handling larger and more diverse datasets, which can affect model generalization. Therefore, this study aims to fill these gaps by investigating transfer learning approaches using ResNet50, both with and without fine-tuning, to enhance the accuracy of pneumonia detection from chest X-ray images.

Based on the aforementioned gaps, the objectives of this research are as follows a) To evaluate the effectiveness of ResNet50 in pneumonia detection from chest X-ray images using transfer learning approaches. b)To compare the performance of ResNet50 with and without fine-tuning in terms of accuracy and other relevant performance metrics. c) To investigate the impact of using larger and more diverse datasets on the generalization capabilities of the ResNet50 model in pneumonia detection. d) To provide insights and recommendations for optimizing transfer learning techniques in deep learning models for medical image analysis, specifically in the context of pneumonia detection.

Thus, this study is expected to contribute significantly to improving the accuracy and efficiency of pneumonia detection, as well as addressing several challenges associated with implementing deep learning technologies in the medical field.

2. MATERIAL AND METHOD

2.1 Type of Research

This study is an experimental quantitative research aimed at comparing two transfer learning approaches using ResNet-50 for pneumonia detection from chest X-ray images: one with fine-tuning and one without fine-tuning (frozen layers). The research was conducted to evaluate the effectiveness and performance of ResNet-50 under these two configurations, utilizing a well-established dataset of chest X-rays. The primary focus is to assess accuracy and other relevant performance metrics and understand how fine-tuning affects the model's generalization.

2.2 Research Object and Sampling Method

The research object consists of chest X-ray images labeled as *Pneumonia* or *Normal*, sourced from the Kaggle Chest X-Ray Pneumonia Dataset (public dataset). This dataset contains 5,856 images, split into two classes: pneumonia and normal.

The stratified sampling method was employed to ensure a balanced distribution of both labels across the training, validation, and test sets. The dataset was divided as follows: 70% for training, 20% for testing, 10% for validation.

Additionally, care was taken to ensure patient-level separation between these sets to prevent data leakage, ensuring that the same patient's data does not appear in both the training and testing sets.

2.3. Research Procedures

2.3.1 Dataset Acquisition

The dataset was obtained from Kaggle and contains labeled images in two main categories: pneumonia and normal. The dataset is publicly available and widely used in pneumonia detection studies.

2.3.2 Data Preprocessing

Image resizing was performed to standardize image dimensions to 224 × 224 pixels.

Normalization was applied to scale pixel values to the range [0–1].

Data augmentation was applied to the training set to improve model robustness, including techniques such as:

a. *Random rotation* (±15°),
b. *Horizontal flip* (probability = 0.5),
c. *Zooming* (0.1),
d. *Shifting* (0.1).

The validation and test sets were not augmented to ensure accurate evaluation of the model's performance.

2.3.3 Model Development

Two configurations of the ResNet-50 model were developed:

a. Transfer Learning – Frozen Layers (No Fine-Tuning). All layers of the ResNet-50 backbone were frozen. Only the fully connected classifier was trained.
b. Transfer Learning – Fine-Tuning (With Fine-Tuning). The last 30 layers of the ResNet-50 were unfrozen, meaning their parameters were updated during training.

2.3.4 Hyperparameters and Training Configuration

a. Optimizer: Adam optimizer was used for training.
b. Learning Rate: 1e-4 for the frozen model, and 5e-5 for the fine-tuned model.
c. Batch size: 32
d. Epochs: Training was conducted over 25–30 epochs, with early stopping applied (patience = 7).
e. To improve model robustness, k-fold cross-validation was implemented for parameter tuning, and stratified sampling was used to ensure a balanced distribution of data in all splits.

2.3.5 Model Evaluation

The models were evaluated using the test set, and performance was measured using the following metrics:

a. Accuracy,
b. Precision,
c. Recall,
d. F1-score,
e. AUC (Area Under the Curve) for ROC curves.
f. These metrics were used to compare the performance of the fine-tuned and frozen models.

2.4. Data Collection Techniques

Data was primarily collected from the Kaggle Chest X-Ray Pneumonia Dataset. During the study, the following output data was collected:

1) Training logs,
2) Loss curves and accuracy curves,
3) Grad-CAM visualizations for model interpretability.

Data was processed using Python scripts and computational tools such as TensorFlow, Keras, and Scikit-learn to calculate and analyze performance metrics.

2.5. Research Instruments

The instruments used in this study include:

Hardware

1) Google Colab Pro with access to NVIDIA Tesla T4/A100 GPUs
2) RAM: 12–25 GB

Software

1) Python 3.10
2) TensorFlow / Keras for model development
3) Matplotlib, NumPy, Pandas for data manipulation and visualization
4) Scikit-learn for performance metrics (Confusion Matrix, AUC)
5) Grad-CAM Toolkit for model interpretability

2.6. Data Analysis Plan

Data analysis was conducted using both computational methods and statistical tests to evaluate the models' performance.

Performance Metrics

Accuracy, Precision, Recall, F1-score, and AUC from ROC curves were used to assess the models.

Statistical Methods

To compare the performance between the two models (frozen vs fine-tuned), paired t-tests or Wilcoxon signed-rank tests were applied to determine the statistical significance of performance differences.

Mean ± Standard Deviation from multiple training runs (using different random seeds) was also reported to assess the variability of results.

2.7 Visualization

1) Confusion Matrix
2) ROC and Precision-Recall curves
3) Learning curves (showing training vs validation performance)

2.8. Scope and Limitations

The scope of this study is focused on comparing two configurations of ResNet-50 for pneumonia detection from chest X-ray images: one with frozen layers and one

with fine-tuning. The study has the following limitations:

1) Single dataset usage: This study relies on one publicly available dataset, which limits generalizability to different institutions or populations.
2) No comparison with other architectures: Other deep learning architectures such as DenseNet, EfficientNet, or Vision Transformers were not included in this study for comparison.
3) Binary classification: The study focuses on binary classification (pneumonia vs normal) and does not include multi-class classification (e.g., bacterial vs viral pneumonia).
4) No evaluation of model inference time: Computational inference time and model parameter count were not assessed, which are important for real-world clinical deployment.
5) Interpretability: While Grad-CAM is used for model interpretability, a more quantitative analysis of the model's interpretability (e.g., overlap ratio with lung segmentation masks) could strengthen the findings.

3. RESULT AND DISCUSSION

3.1 Result

3.1.2 Performance Metrics

Frozen Model Metrics:
a. accuracy: 0.6250
b. precision: 0.6250
c. recall: 1.0000
d. f1: 0.7692
e. auc: 0.4819

Fine-tuned Model Metrics:
a. accuracy: 0.8590
b. precision: 0.8268
c. recall: 0.9795
d. f1: 0.8967
e. auc: 0.9510

The fine-tuned model clearly outperforms the frozen version across nearly all evaluation metrics. Its AUC rises from 0.48 to 0.95, indicating a much stronger ability to separate the two classes correctly. After fine-tuning, the model becomes more reliable overall, showing higher accuracy, better precision, and a more balanced performance.
Although recall drops slightly, it remains extremely high, meaning the model still identifies almost all positive samples.

3.1.3 Statistical Methods (per-fold)

Accuracy: mean(finetune)=0.8141
mean(frozen)=0.6250
Paired t-test: t=7.7920, p=0.016074
Wilcoxon: W=0.0000, p=0.250000
AUC:
mean(finetune)=0.9528, mean(frozen)=0.6301
Paired t-test: t=4.1262, p=0.054021
Wilcoxon: W=0.0000, p=0.250000

3.1.4 Overall Summary

a. For accuracy, the fine-tuned model clearly performs better, and this improvement is statistically supported by the paired t-test.
b. For AUC, the fine-tuned model also shows a large practical improvement, but the statistical tests do not provide strong evidence of significance—likely due to the limited number of folds.
c. In practical terms, the fine-tuned model offers consistently better performance, even though not all statistical tests capture the difference.

3.1.5 Visualization

a. Confusion Matrix

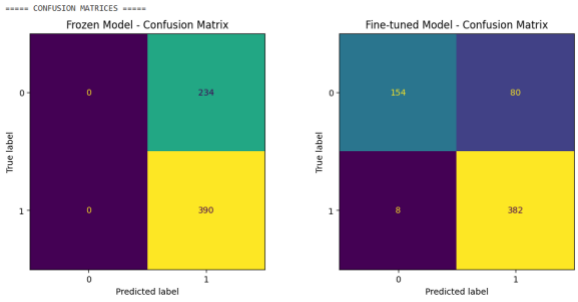

Figure 1. Confusion matrix of ResNet-50 Frozen and Fine Tuning

The confusion matrices provide a clear comparison of how the frozen and fine-tuned models behave on the classification task. In the frozen model, all instances of class 0 are misclassified as class 1, indicating that the model entirely fails to recognize one of the classes. Although it correctly labels every sample belonging to class 1, the inability to detect class 0 results in a highly imbalanced and unreliable classifier.

In contrast, the fine-tuned model demonstrates substantial improvement. It correctly identifies a large portion of class-0 samples while maintaining strong performance on class 1, with only a small number of errors in each category. This shift shows that fine-tuning enables the model to develop a more balanced understanding of the data, leading to a classifier that is more robust, more discriminative, and considerably more effective than the frozen version.

b. ROC and Precision-Recall curves

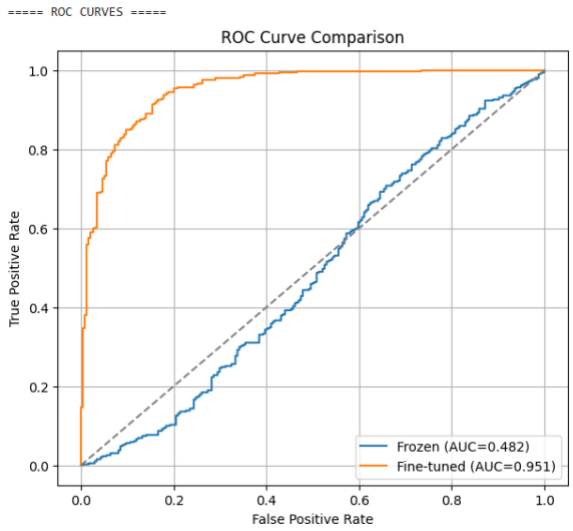

Figure 2. ROC Curve Comparison

The ROC curve highlights a substantial performance gap between the frozen model and the fine-tuned model. The curve for the frozen model lies close to the diagonal reference line, which indicates that its predictions are only slightly better than random guessing. This is further supported by its low AUC value of 0.482, showing that the model has very limited ability to distinguish between the two classes. In contrast, the fine-tuned model produces a ROC curve that bends sharply toward the upper-left corner, reflecting a strong balance between high true positive rates and low false positive rates. Its AUC score of 0.951 demonstrates a highly effective classifier capable of separating the classes with a high level of reliability. Overall, the ROC comparison clearly shows that fine-tuning dramatically enhances the model's discriminative power.

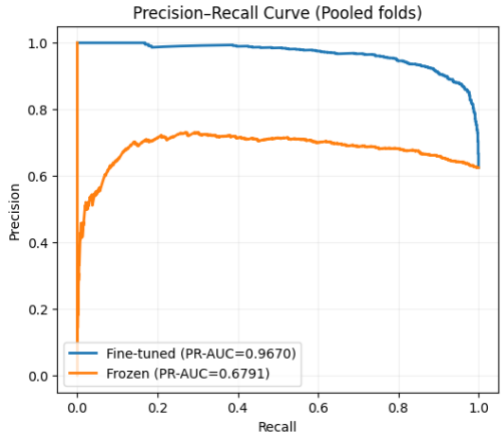

Figure 3. Precision Recall Curve

3.2 Discussion

This study demonstrates that fine-tuning significantly improves the diagnostic performance of the ResNet-50 architecture for pneumonia detection using chest X-ray images. The fine-tuned model achieved an accuracy of 85.90%, F1-score of 0.8967, and AUC of 0.951, while the frozen model produced substantially weaker results, with accuracy of 62.50% and AUC of only 0.4819.

This performance gap indicates that freezing convolutional layers prevents the model from adapting feature representations to pneumonia-specific patterns, such as pulmonary opacity and interstitial markings. Since ResNet-50 is pre-trained on ImageNet (natural images), its higher-level features do not align naturally with radiographic characteristics unless fine-tuning is applied.

These results are consistent with previous research emphasizing that domain adaptation significantly enhances model reliability in medical imaging (Musa et al., 2024, 2025). Furthermore, the fine-tuned model maintained both high recall (97.95%) and high precision (82.68%), demonstrating a robust balance between identifying infected patients and minimizing false alarms.

The confusion matrix reveals critical behavioral differences between the two models. The frozen model classified nearly all samples as pneumonia, which artificially inflated recall but severely compromised specificity. This results in a biased classifier that fails to distinguish between diseased and healthy images.

In contrast, the fine-tuned model demonstrated a more balanced distribution of predictions across both classes. This indicates that fine-tuning allows deeper convolutional layers to recalibrate their learned filters toward pathology-driven features rather than relying on general textures.

From a clinical viewpoint, the frozen model's high false-positive rate would cause unnecessary clinical follow-up and increased hospital burden, whereas the fine-tuned model reduces both false positives and false negatives. Similar concerns regarding diagnostic imbalance have been emphasized in previous medical imaging studies (Banik & Bhattacharjee, 2022; Cabeza et al., 2025; Rguibi et al., 2022).

The ROC curve further illustrates the disparity between the two approaches. The frozen configuration exhibited an ROC curve close to the diagonal reference line, corresponding to an AUC of **0.4819**—suggesting near-random classification performance.

The fine-tuned model, conversely, showed a steep ROC curve approaching the upper-left corner, with an outstanding AUC of **0.951**, indicating strong discriminatory capability between pneumonia and normal cases.

Precision–Recall curves reinforced this conclusion. The frozen model sacrificed precision in favor of recall, leading to an unstable predictive pattern. The fine-tuned model maintained consistently higher precision across recall levels, reflecting clinical reliability under varying diagnostic thresholds.

3.2.1 Implications

The findings of this study highlight several important implications for both research and clinical practice. From a methodological standpoint, the results demonstrate that fine-tuning is a necessary strategy when adapting deep learning models trained on natural images to medical imaging tasks. Models that rely solely on frozen feature extractors are likely to suffer from representation mismatch, which significantly reduces classification performance.

In clinical contexts, the superior performance of the fine-tuned model suggests that properly adapted CNN architectures can assist healthcare professionals in early pneumonia screening. This is particularly relevant in settings where access to radiologists is limited. Automated systems based on deep learning can be used as supportive tools to enhance diagnostic efficiency, reduce delays, and improve overall patient triage processes. Nevertheless, these systems should function as assistive technologies rather than replacements for clinical judgment.

3.2.2 Research contribution

This research contributes to the existing body of knowledge in several ways. First, it offers a direct empirical comparison between frozen and fine-tuned transfer learning strategies using the same architecture and dataset, providing clear evidence of their performance differences in pneumonia detection. Unlike many previous studies that only report model accuracy, this study emphasizes multiple performance metrics, including AUC, F1-score, and confusion matrix patterns.

Second, the study illustrates that freezing convolutional layers may lead to severe classification bias, where the model becomes incapable of recognizing one of the target classes. This highlights a rarely discussed failure mode in transfer learning applications. Finally, this research reinforces the theoretical understanding of domain adaptation in deep learning and demonstrates its importance in medical imaging applications.

3.2.3 Limitations

Despite the promising results, this study has several limitations. First, the dataset used was limited to one public source, which restricts the generalizability of the findings across different healthcare institutions and populations. Second, the classification task was limited to a binary scenario (pneumonia and normal), which does not reflect the complexity of real-world diagnosis involving multiple pneumonia types. Third, the evaluation did not include real-time performance metrics such as inference speed or computational resource usage.

Additionally, although interpretation tools were used, the study did not perform quantitative evaluation of explanation quality. Lastly, although k-fold cross-validation was applied, the small number of folds may have limited the statistical power of hypothesis testing.

3.2.4 Suggestions

Future research should address the limitations identified in this study. Multi-center datasets should be incorporated to evaluate model robustness across different imaging environments and patient populations. Further investigations should also explore multi-class classification approaches to distinguish between viral, bacterial, and other lung conditions.

Comparative studies involving newer architectures such as EfficientNet, DenseNet, and Vision Transformers may provide additional insights into optimization strategies. Moreover, future work should evaluate computational efficiency to ensure deployment feasibility in real clinical systems. Lastly, combining chest X-ray data with clinical variables such as age, symptoms, and oxygen saturation may improve decision-making accuracy.

4. CONCLUSION

This study demonstrates that transfer learning with fine-tuning significantly improves the performance of ResNet-50 in pneumonia classification using chest X-ray images. While the frozen-layer model failed to generalize effectively and exhibited biased classification behavior, the fine-tuned model achieved strong diagnostic performance, as reflected by high accuracy, AUC, and F1-score.

The results confirm that domain adaptation is essential when applying pretrained neural networks to medical imaging tasks. Fine-tuning enables the network to learn disease-specific features rather than relying on irrelevant characteristics acquired from natural images. Consequently, the fine-tuned ResNet-50 model provides a more reliable foundation for computer-aided pneumonia detection.

In conclusion, fine-tuning should be considered a fundamental component in transfer learning pipelines for medical imaging. The findings of this study support further investigation into the development of clinically usable deep learning models using larger datasets and broader evaluation frameworks.

5. ACKNOWLEDGEMENT

The authors would like to sincerely thank Politeknik Pajajaran and Institut Kesehatan Immanuel for their continuous support and encouragement throughout the completion of this research. The availability of institutional resources and academic guidance greatly assisted the authors in conducting the study. We also extend our appreciation to our colleagues and peers for their constructive input and motivation, which significantly enhanced the overall quality of this work.

AUTHOR INFORMATION

Corresponding Authors

Dwi Robiul Rochmawati, Politeknik Pajajaran, Bandung, Indonesia
ORCiD: https://orcid.org/0009-0008-1595-2430
Email: dwirobiul@gmail.com

Authors

Lidya Maryani, Faculty of Nursing, Institut Kesehatan Immanuel, Bandung, Indonesia
ORCiD: https://orcid.org/0000-0001-8825-7940
Email: lidyashare@gmail.com

REFERENCE

Aditya Pai, H., Pareek, P. K., Guruprasad, M. S., Chandrappa, S., & Bhargava, K. (2024). *Pneumonia Detection in Chest X-Rays Using Combined Approach of Deep Learning*. 2nd IEEE International Conference on Networks, Multimedia and Information Technology, NMITCON 2024. Scopus. https://doi.org/10.1109/NMITCON62075.2024.10699169

Alshanketi, F., Alharbi, A., Kuruvilla, M., Mahzoon, V., Siddiqui, S. T., Rana, N., & Tahir, A. (2025). Pneumonia Detection from Chest X-Ray Images Using Deep Learning and Transfer Learning for Imbalanced Datasets. *Journal of Imaging Informatics in Medicine*, *38*(4), 2021–2040. Scopus. https://doi.org/10.1007/s10278-024-01334-0

Banik, D., & Bhattacharjee, D. (2022). Mitigating data imbalance issues in medical image analysis. In *Research Anthology on Improving Medical Imaging Techniques for Analysis and Intervention* (pp. 1215–1238). Scopus. https://doi.org/10.4018/978-1-6684-7544-7.ch063

Bhola, A., & Gulhane, M. (2024). Revolutionizing Pneumonia Diagnosis and Prediction Through Deep Neural Networks. In *Optimized Predictive Models in Health Care Using Machine Learning* (pp. 135–150). Scopus. https://doi.org/10.1002/9781394175376.ch9

Cabeza, N., Fajardo, C. A., & Pertuz, S. (2025). *Evaluating the Impact of Class Imbalance on Breast Ultrasound Image Classification*. 2025 25th Symposium of Image, Signal Processing, and Artificial Vision, STSIVA 2025. Scopus. https://doi.org/10.1109/STSIVA66383.2025.11156656

Dabre, K., Varma, S. L., & Patil, P. B. (2024). RAPID-Net: Reduced architecture for pneumonia in infants detection using deep convolutional framework using chest radiograph. *Biomedical Signal Processing and Control*, *87*. Scopus. https://doi.org/10.1016/j.bspc.2023.105375

Gupta, A., Arora, S., Jain, M., & Jain, K. (2025). *Comparative Analysis of ResNet-18 and ResNet-50 Architectures for Pneumonia Detection in Medical Imaging*. *1096 LNNS*, 355–365. Scopus. https://doi.org/10.1007/978-981-97-7178-3_31

Jihad, M., Siam, S., Anannya, L. A., Anonto, F. R., Safin Kamal, K. M., & Reza, A. W. (2025). *Explainable Artificial Intelligence for Deep Learning-Based Detection of Pneumonia in Chest X-ray Images*. *1408 LNNS*, 273–287. Scopus. https://doi.org/10.1007/978-981-96-6297-5_21

Khojare, S., & Bhattacharjee, S. (2025). *Pneumonia Identification based on Transfer Learning over Chest X-Ray Images*. Proceedings of 2025 International Conference on Emerging Trends in Industry 4.0 Technologies, ICETI4T 2025. Scopus. https://doi.org/10.1109/ICETI4T63625.2025.11132185

Maheshwar, A. M., Hari Prasad, S., & Hemalatha, R. (2025). *Pneumonia Detection using Chest X-Rays*. 1396–1402. Scopus. https://doi.org/10.1109/ICICCS65191.2025.10984571

Manickam, A., Jiang, J., Zhou, Y., Sagar, A., Soundrapandiyan, R., & Dinesh Jackson Samuel, R. (2021). Automated pneumonia detection on chest X-ray images: A deep learning approach with different optimizers and transfer learning architectures. *Measurement: Journal of the International Measurement Confederation*, *184*. Scopus. https://doi.org/10.1016/j.measurement.2021.109953

Michele, A., & Kusuma, G. P. (2021). BACTERIAL and VIRUS PNEUMONIA INFECTION DETECTION on CHEST X-RAY IMAGES USING MACHINE LEARNING. *Journal of Theoretical and Applied Information Technology*, *99*(24), 5638–5649. Scopus.

Munna, M. S., Plabon, N. S., & Emon, A. P. (2025). *Real-Time Multi-Class Pneumonia Detection in Chest X-Rays Using Optimized Deep Learning Methods*. 2025 IEEE International Conference on Quantum Photonics, Artificial Intelligence, and Networking, QPAIN 2025. Scopus. https://doi.org/10.1109/QPAIN66474.2025.11171805

Musa, A., Adamu, M. I., Kakudi, H. A., Hernandez, M., & Lawal, Y. (2024). *Analyzing Cross-Population Domain Shift in Chest X-Ray Image Classification and Mitigating the Gap with Deep Supervised Domain Adaptation*. *15003 LNCS*, 585–595. Scopus. https://doi.org/10.1007/978-3-031-72384-1_55

Musa, A., Prasad, R., & Hernandez, M. (2025). Addressing cross-population domain shift in

chest X-ray classification through supervised adversarial domain adaptation. *Scientific Reports*, *15*(1). Scopus. https://doi.org/10.1038/s41598-025-95390-3

Paul, M., & Naskar, R. (2023). *Deep Learning Enabled Pneumonia Detection From Chest X-rays: A Transfer Learning Based Ensemble Classification Approach*. 2023 IEEE 3rd International Conference on Smart Technologies for Power, Energy and Control, STPEC 2023. Scopus. https://doi.org/10.1109/STPEC59253.2023.10430698

Rani, S., & Kaur, N. (2024). *Comparative Analysis of Pneumonia Detection from Chest X-ray Using Deep Learning*. *1001 LNNS*, 276–287. Scopus. https://doi.org/10.1007/978-3-031-60935-0_25

Rguibi, Z., Hajami, A., & Dya, Z. (2022). Automatic Searching of Deep Neural Networks for Medical Imaging Diagnostic. In *Lecture Notes on Data Engineering and Communications Technologies* (Vol. 110, pp. 129–140). Scopus. https://doi.org/10.1007/978-3-030-94188-8_13

Shah, U., Abd-Alrazeq, A., Alam, T., Househ, M., & Shah, Z. (2020a). *An efficient method to predict pneumonia from chest X-rays using deep learning approach*. *272*, 457–460. Scopus. https://doi.org/10.3233/SHTI200594

Shah, U., Abd-Alrazeq, A., Alam, T., Househ, M., & Shah, Z. (2020b). *An efficient method to predict pneumonia from chest X-rays using deep learning approach*. *272*, 457–460. Scopus. https://doi.org/10.3233/SHTI200594

Shakeri, E., & Far, B. (2024a). *Advancing Pneumonia Classification and Detection: Comparative Analysis of Deep Learning Models Using Convolutional Neural Networks*. 108–113. Scopus. https://doi.org/10.1109/IRI62200.2024.00033

Shakeri, E., & Far, B. (2024b). *Advancing Pneumonia Classification and Detection: Comparative Analysis of Deep Learning Models Using Convolutional Neural Networks*. 108–113. Scopus. https://doi.org/10.1109/IRI62200.2024.00033

Sharma, S., & Guleria, K. (2024). A systematic literature review on deep learning approaches for pneumonia detection using chest X-ray images. *Multimedia Tools and Applications*, *83*(8), 24101–24151. Scopus. https://doi.org/10.1007/s11042-023-16419-1

Shirvalkar, R., & Remya Ajai, A. S. (2024). *Enhancing Pneumonia Detection Accuracy Through ResNet-Based Deep Learning Models and Ensemble Techniques: A Study Using Chest X-Ray Images*. *949 LNNS*, 25–37. Scopus. https://doi.org/10.1007/978-981-97-1313-4_3

Shrimali, S. (2024). *Optimizing Chest X-ray Analysis for Pneumonia Detection Through Comparative Evaluation of Transfer Learning CNN Architectures*. 2024 IEEE 9th International Conference for Convergence in Technology, I2CT 2024. Scopus. https://doi.org/10.1109/I2CT61223.2024.10543957

Sohel, A., Rahman, M. M., Zubaer, T., Rifat, R. I., Majumder, P. P., & Bijoy, M. H. I. (2024). *Improving Pneumonia Detection in X-ray Images with Hybrid Deep Learning Techniques*. 2024 6th International Conference on Sustainable Technologies for Industry 5.0, STI 2024. Scopus. https://doi.org/10.1109/STI64222.2024.10951093

Surani, S., & Jain, N. K. (2025). *Bronchopneumonia: Pathogenesis, clinical insights and therapeutic approaches* (p. 492). Scopus. https://doi.org/10.52305/EVQE1662

Vemuri, S. S., Kotha, S., Voruganti, S., Kunam, P. R., Nandini, A., & Kumar Mishra, T. (2025). *Pneumonia Detection from X-Ray Images Using Deep Transfer Learning*. *2317 CCIS*, 205–219. Scopus. https://doi.org/10.1007/978-3-031-79041-6_17

Virmani, A., Singh, A., Agarwal, R., Kumar, S., & Kumar, H. (2023). X-Ray-Based Pneumonia Detection Using ResNet50 and VGG16 Extracted Features and Conventional Machine Learning Algorithms. In *Artificial Intelligence and Machine Learning: An Intelligent Perspective of Emerging Technologies* (pp. 15–25). Scopus. https://doi.org/10.1201/9781003388319-2

Wang, K., Zhang, X., Huang, S., & Chen, F. (2019). *Automatic detection of pneumonia in chest X-ray images using cooperative convolutional neural networks*. *11858 LNCS*, 328–340. Scopus. https://doi.org/10.1007/978-3-030-31723-2_28

Yaqin, A., Lintang Pamungkas, O. T., Satya, B., Kraugusteeliana, K., Frobenius, A. C., & Ferdiansyah, P. (2024). *Pneumonia Disease Analysis Using Deep Learning on X-Ray Images*. 804–808. Scopus. https://doi.org/10.1109/ICIMCIS63449.2024.10956956